# Open Data, Information Processing and Datification of Government

Marijn Janssen
Delft University of Technology
m.f.w.h.a.janssen@tudelft.nl

Yannis Charalabidis
University of the Aegean
yannisx@aegean.gr

Helmut Krcmar
Technische Universität München
Krcmar@in.tum.de

## Abstract

*Collecting, combining and processing all kinds of data is a key element driving public sector transformation. A conditions for this is interoperability, referring to the ability of diverse systems to work together, and infrastructures facilitating these activities. In the past, the focus of this minitrack was on infrastructure, which have changed over time. Over time infrastructures, clouds and open data have become more important. Current developments results in a datafication of government. This development is driven by the collecting of more and more data using the Internet of Things (IoT) and the sharing and opening of data. This results in the need for (big) data information sharing and processing infrastructure and capabilities. All these developments influence many aspects of government ranging from fraud detection to policy-making. Administrative organizations become largely driven by ICT and the ability to manage ICT has become key to their successful functioning.*

## 1. Introduction

The public sector is data-intensive by its very nature. Developments like the opening of data and the Internet of Things (IoT) results in the availability of more data in digital format, which can be used by organizations to improve their decision- and policy-making. Datification refers to sensing and the subsequent collecting of all kinds of data in machine-readable data formats [1]. Datafication is rapidly becoming a mainstream activity of public organizations and can be used in various areas.

Another development is the opening of government data to the public. Businesses and citizens can combine open data with all kinds of data sources to infer and generate value. This can results in benefits like recommendations for improving the public sector, business model innovation, and the creation of transparency.

The rise of all kinds of data has resulted in the demand for new approaches for organizing, storing, processing, curation, linking and visualization the result of data. Data pipelines are created in which data is combined in real-time for creating new applications. Cloud services are changing the ways of providing and processing data, based on virtualized resources meeting requirements like security, privacy and scalability [2]. The development of infrastructures have become a key component of many governments to facilitate reuse and development of new applications. New applications are often development in an agile manner [3] and organizations need to be adaptive [4].

All these developments are resulting in drastic changes of the public sector. Although there is a huge potential, how this should be accomplished, and the impact on public organizations is ill-understood. Furthermore there is a need to deal with constraints like privacy [5], often originating from the legislative environment. The operation of governments, their relationship with the public, enforcement activities and policy-making activities are changing at the technical, organizational, managerial and political level. This minitrack is aimed at discussing theories, methodologies, experiences, literature and case studies in the field of open data, information processing and datification in Government. This minitrack is positioned at the intersection of organizational and technical aspects.

## 2. Key contributions

E-government is a multidisciplinary research area and theories and concepts from disciplines like computer science, organizational science, information science, political science, psychology, and public administration contribute to this field. The combination of disciplines makes e-government research often challenging as the value often originates from combining theories and insights originating from different disciplines, which requires both knowledge of the disciplines and deep insight in the e-government domain.

This e-government minitrack was introduced at HICSS39 held in 2006 and the initial name was “e-Gov infrastructure”. This name refers to the need to developed shared infrastructure to connect organizations and to facilitate information sharing.

For the HICSS42 conference held in 2009 the name was changed into “e-Gov Infrastructure and

URI: http://hdl.handle.net/10125/41478
ISBN: 978-0-9981331-0-2

Interoperability". *Interoperability* is the ability of diverse systems to work with each other [6]. Interoperability captures aspects ranging from the technical to the organizational level.

In 2011 (HICSS44) the minitrack name was changed into "Interoperability and cloud infrastructure". This adapted name refers to the emerging field to the use of clouds [2]. Clouds refer to the remote access of virtualized computing power and memory over a communication network, like the Internet. Cloud enables application to run remotely and have changed the way systems are developed, operated and maintained. A small variation o was introduced for HICSS46 in 2013 resulting into the name "Cloud Services and Interoperability".

With the advent of the open data stream the name was changed into "open data and cloud services" for the HICSS47 conference held in 2014. More and more governments started to open their data for use by others, which has resulted in new applications and innovations and is even influencing how governments operate. The opening of data can used to create societal value, but practice lags behind reality.

The next year the complete range of topics was included in the minitrack name "Big, Open, Linked Data (BOLD), Analytics, and Interoperability Infrastructures in Government". The next year this name was simplified by using only "Big, Open, and Linked Data (BOLD)" to stress the importance of data. This was continued this year for HICSS50 as data is at the very heart of many public organizations processes and activities.

The bottom line of the minitrack is the leveraging technology, data, and innovative ideas to improve and transform government. Yet the effects of technology might not always be positive and there might be a dark side. To tackle these issues, often a variety of theoretical lenses need to be adopted. Technology becomes institutionalized after some time. To reap the benefits there is often a need for organizational adoption and a need for transforming organizational structures and practices. Furthermore, governments need to develop new capabilities and develop policies to deal with these developments.

How value can be created is often not clear and typically is one of the key research questions when new technology is used. Addressing this requires understanding of the stakeholders, organizational aspects and the potential of technology.

Papers taking interdisciplinary approaches and covering a multitude of aspects are often received and accepted by this minitrack. The papers employ a diversity of research methods to study the challenges of this multifaceted problem.

## 3. Papers overview

This year 7 papers were accepted resulting in 2 sessions. The first session has 4 papers and starts with the paper authored by Mahdi M. Najafabadi and Luis Luna-Reyes. The paper "*Open Government Data Ecosystems: A Closed-Loop Perspective*" presents a modeling and simulation model of an open data ecosystem. The model can be used to improve the understanding of enablers and barriers of open data and in this way resulting in improvements in open data policy making.

In the second paper of the first session "*Data Collaboratives as a New Frontier of Cross-Sector Partnerships in the Age of Open Data: Taxonomy Development*" Iryna Susha, Marijn Janssen, and Stefaan Verhulst analyze new organizational forms for creating value from data. In their paper a new type of business model is analyzed in which private and public organization collaborate to tackle societal challenges.

Predictive analytics can be used in policy-making. Akemi Takeoka Chatfield and Christopher Reddick analyze the strategic decision-making effectiveness of predictive analytics in their paper "*Barriers to Predictive Analytics Use for Policy Decision-Making Effectiveness in Turbulent Times: A Case Study of Fukushima Nuclear Accident*". Barriers ranging from the institutional to the technical level were found.

Challenges and the expectations of civic hackers are discussed in the paper named "*Civic Hackers' User Experiences and Expectations of Seattle's Open Municipal Data Program*". This paper authored by Meg Young and An Yan reveal that hackers prefer higher quality data and usable data portals having more functionality. The authors plea for a third party to host cleaned data for wider use.

The second session contains 3 papers. The first paper "*The Role of Information Quality in Healthcare Organizations: A Multi-Disciplinary Literature Review*" Geir Inge Hausvik discusses the quality of the exponentially growing volume of data. The authors plea for understanding the implications of (a lack of) information quality to create secure and safe health services.

The second paper of the second session "*Is Organic Labelling Enough? Information Disclosure as Policy Instrument to Empower Consumer Choices*" is authored by Jing Zhang and Lin Boldt. The authors discuss the need to develop measures that help decide what information should be disclosed and whether a disclosure should be mandated. Their analysis suggest a significant difference in consumers' choice when information is disclosed.

The last paper "*Understanding Datafication Effects of Open Government Information Systems - A Contemporary Systems Thinking Approach*" written by Olivera Marjanovic and Dubravka Cecez-Kecmanovic investigates the negative effects of datafication. The researchers found that the dematerialization, liquification and strategizing mechanisms might result in harmful effect. Publishing performance data might be far reaching and the effects difficult to predict.

The 7 papers take a variety of approaches and interdisciplinary views on this domain. They show the developments and changes that are going on, and a variety of implications.

## 4. Further research directions

The many technology development influences many aspects of the public sector. There are many future avenues, including understand value creating mechanisms, new ways of collaborating, institutionalizing open data, developing measurement instruments, and understanding the broader impact. There is a clear need for further theorizing in this domain which requires theories and concepts from multiple disciplines.

There are many developments impacting the government that are ill-understood. The continued evolution of ICT requires research about what the impact is on governments and their relationship with the public. In particular how public value can be created using technologies. The adoption of technology transforms governments in the way they operate and deal with the public. Research must not only address the relationship between ICT and organizational processes but should understand how these changes influence public value generation and society as a whole. Finally, organizational practices and procedures need to be adopted and institutionalized.

Government activities have become largely driven by ICT. The ability to manage and governance of ICT has become key to their successful functioning and needs ample further research. Infrastructures need to be developed that facilitate the collecting, processing and use of data originating from many sources. This data is used in an ecosystem of actors. How the infrastructure and ecosystem can be developed and looks like is unclear. Parts owned by public and private parties can be used, which demands coordination and governance.

Policy-making quality is influenced by the data and the process to make sense of the data [4]. This process should be analyzed to better understand its consequences on the decisions and policies that are based on this data.

Open government needs collaboration and knowledge sharing amongst. Organizations cannot act in isolation and have to act within an ecosystem. This requires awareness of other players and the need for investigating capabilities that can be used by the organizations to operate in such an ecosystem.

Datafication is driving the use of *algorithms* for decision-making, which might have unintended consequences [7]. The inclusion of artificial intelligence in systems will likely be one of the major next challenges in this field. We are on the eve of this development and this will likely trigger the need for more research.

Finally, all these developments have a profound on aspects like openness, transparency, accountability, engagement, security and anti-corruption, that are essential public values of governments. How these values are affected and how we can realize this values needs to be unraveled.

## 5. References


[1] V. Mayer-Schönberger, and K. Cukier, *Big data: A revolution that will transform how we live, work, and think*: Houghton Mifflin Harcourt, 2013.

[2] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandi, "Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility," *Future Generation Computer Systems*, vol. 25, no. 6, pp. 599-616, 2009.

[3] I. Mergel, "Agile innovation management in government: A research agenda," *Government Information Quarterly*, vol. 33, no. 3, pp. 516-523, 7//, 2016.

[4] M. Jansen, H. Van der Voort, and A. Wahyudi, "Factors influencing big data decision-making quality," *Journal of Business Research*, 2016.

[5] L. van Zoonen, "Privacy concerns in smart cities," *Government Information Quarterly*.

[6] H. J. Scholl, and R. Klischewski, "E-Government Integration and Interoperability: Framing the Research Agenda," *International Journal of Public Administration (IJPA)*, vol. 30, no. 8-9, pp. 889-920, 2007.

[7] M. Janssen, and G. Kuk, "The challenges and limits of big data algorithms in technocratic governance," *Government Information Quarterly*, vol. 33, no. 3, pp. 371-377, 2016.